**Loneliness - What characteristics and circumstances are associated with feeling lonely?**

Ed Pyle and Dani Evans; Office for National Statistics

**Abstract**

Loneliness can seriously damage health and well-being ([Holt-Lunstad, 2015](http://journals.sagepub.com/doi/abs/10.1177/1745691614568352)). Though it can affect those of all ages, from different backgrounds, we know that individuals with particular characteristics and circumstances are more vulnerable. In January, the [Prime Minster](https://www.gov.uk/government/news/pm-commits-to-government-wide-drive-to-tackle-loneliness) announced the development of a strategy to alleviate loneliness and requested that ONS develop national measures of loneliness. As part of this, we published [Loneliness – What characteristics and circumstances are associated with feeling lonely?](https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/lonelinesswhatcharacteristicsandcircumstancesareassociatedwithfeelinglonely/2018-04-10) using measures and data already available.

The Community Life Survey 2016-17 ([DCMS, 2017](https://www.gov.uk/government/collections/community-life-survey)) was used for this analysis. Initially descriptive analysis was carried out to shed light on how personal characteristics and circumstances are associated with self-reported loneliness. Logistic regression and LCA analysis approach the exploration of loneliness from two different, but complementary, standpoints. The logistic regression was used to isolate single factors associated with loneliness. The LCA was used to identify combinations of factors that frequently appear together among those who report feeling lonely more often. LCA helps to provide a more holistic picture than logistic regression, highlighting that, in reality, combinations of characteristics tend to go together in ways that influence feelings of loneliness.

LCA was used to classify people experiencing loneliness most often into groups of people with similar characteristics and life-circumstances. The analysis produced different loneliness ‘profiles’ – combinations of characteristics that group together in ways associated with loneliness – and allowed us to identify particularly lonely people and find out more about them. Three profiles were identified: (i) Widowed older homeowners living alone with long-term health conditions; (ii) Unmarried, middle-agers with long-term health conditions; and (iii) Younger renters with little trust and sense of belonging to their area. This information helps to understand the types of people more at risk of feeling lonely and the potential explanations as to why people feel lonely; this has been used for informing the strategy and in turn future policies and interventions aimed at ameliorating loneliness.

**Keywords**: Loneliness; latent class analysis; binomial logistic regression; well-being.

Corresponding author: Ed Pyle

Telephone: 01329 44 7141

Email: edward.pyle@ons.gov.uk

# **Main points**

* In 2016 to 2017, there were 5% of adults in England who reported feeling lonely “often” or “always”.
* Younger adults aged 16 to 24 years reported feeling lonely more often than those in older age groups.
* Women reported feeling lonely more often than men.
* Those single or widowed were at particular risk of experiencing loneliness more often.
* People in poor health or who have conditions they describe as “limiting” were also at particular risk of feeling lonely more often.
* Renters reported feeling lonely more often than homeowners.
* People who feel that they belong less strongly to their neighbourhood reported feeling lonely more often.
* People who have little trust of others in their local area reported feeling lonely more often.

Three profiles of people at particular risk from loneliness were identified:

* Widowed older homeowners living alone with long-term health conditions.
* Unmarried, middle-agers with long-term health conditions.
* Younger renters with little trust and sense of belonging to their area.

# **Things you need to know**

The Prime Minister recently announced [the development of a strategy to alleviate loneliness](https://www.gov.uk/government/news/pm-commits-to-government-wide-drive-to-tackle-loneliness) in response to the [report of the Jo Cox Commission on Loneliness (PDF, 2.56MB)](https://www.jocoxloneliness.org/pdf/a_call_to_action.pdf) published in December 2017. As part of this, she requested that Office for National Statistics (ONS) develops national measures of loneliness. ONS is now working with a cross-government group, charities, academics and other stakeholders to review the measurement of loneliness and publish recommendations on this later this year.

This paper is based on data collected for the [Community Life Survey (CLS)](https://www.gov.uk/government/collections/community-life-survey) between August 2016 and March 2017. The survey asked people living in England: “How often do you feel lonely?” with the following response categories: “often/always”, “sometimes”, “occasionally”, “hardly ever” and “never”. The analysis presented here is based on this question. The survey was selected for analysis because it asked respondents about how often they experience loneliness and about other aspects of their lives – this paper is focused on the frequency of loneliness rather than degree of loneliness.

The focus of this paper is on identifying personal characteristics and circumstances that increase or reduce the likelihood of experiencing loneliness. We also present profiles of loneliness – collections of personal characteristics that can put people at greater or lesser risk of loneliness. This could be used to target support more effectively towards people who are at greater risk of feeling lonely more often.

To improve statistical quality, for the logistic regression and the cluster analysis, those who reported feeling lonely “often/always”, “sometimes” or “occasionally” are classified as “more often lonely” whilst those who reported “hardly ever” or “never” feeling lonely are classified as “hardly ever or never lonely”.

The statistical significance of differences noted within this paper are based on non-overlapping 95% confidence intervals.

# **Who is lonely more often?**

Initially descriptive analysis was carried out to shed light on how personal characteristics and circumstances are associated with self-reported loneliness.

Figure 1 shows that in 2016 to 2017, there were 5% of adults (aged 16 years and over) in England reporting feeling lonely “often/always” – that’s 1 in 20 adults. Furthermore, 16% of adults reported feeling lonely sometimes and 24% occasionally.

****

**Women report experiencing loneliness more often than men**

Women reported feeling lonely more frequently than men. They were significantly more likely than men to report feeling lonely “often/always”, “some of the time” and “occasionally” and were much less likely than men to say they “never” felt lonely (Figure 2).

It is possible that this may reflect in part differences in how men and women reflect on their personal experiences of loneliness or respond to the question. Some research suggests that men may be more reluctant than women to report undesirable feelings such as loneliness ([Borys and Perlman, 1985](http://journals.sagepub.com/doi/10.1177/0146167285111006); [Nicolaisen and Thorsen, 2014](http://journals.sagepub.com/doi/10.2190/AG.78.3.b)).



**Younger adults report experiencing loneliness more often**

Compared with all other age groups except the 25 to 34 years group, those aged 16 to 24 years were significantly more likely to report feeling lonely “often/always”. Those aged 16 to 24 years were also significantly more likely to report feeling lonely “some of the time” compared with all other groups except for the 25 to 34 years and 75 years and over age groups. They were also least likely of all age groups to report “never” experiencing loneliness (Figure 3).



**Those widowed report experiencing loneliness more often**

Those who were widowed were significantly more likely to report feeling lonely “some of the time” and least likely to report “never” experiencing loneliness compared with other marital groups.

By contrast, people who were married or in a civil partnership were significantly less likely to report experiencing loneliness “often/always”, “some of the time” or “occasionally”. Consistent with this, those married or in a civil partnership were found to be significantly more likely to report feeling lonely “hardly ever” and “never” (Figure 4).



**Notes:**

1. Married or civil partnership does not include those who reported being separated.
2. Separated includes people who are legally married but reported being separated.

**Those with poorer health report experiencing loneliness more often**

Those who reported their general health to be “very bad” or “bad” were significantly more likely to report feeling lonely “often/always” and significantly less likely to say they “hardly ever” felt lonely compared with all other groups.

People who said their general health was “very good” or “good” were significantly less likely to report feeling lonely “often/always” or “some of the time” and significantly more likely to report “hardly ever” or “never” feeling lonely compared with all other groups (Figure 5).



**Those with a long-term illness or disability report experiencing loneliness more often**

Those who reported having a long-term illness or disability were significantly more likely to report feeling lonely “often/always” and “some of the time”. Those who did not have an illness or disability were significantly more likely to say that they “hardly ever” or “never” felt lonely (Figure 6).



**Unemployed people report being lonely more often**

People who were unemployed (and seeking work) were significantly more likely to report loneliness “often/always” than those in employment or self-employment. Those in employment were significantly more likely to say they “hardly ever” feel lonely than those who were economically inactive (Figure 7). People who are considered “economically inactive” include those not in employment or seeking work, including retirees.



**People living as a couple are lonely less often**

People who were not living as part of a couple were significantly more likely to report experiencing loneliness “often/always”, “some of the time” or “occasionally” than those who reported living as part of a couple. People who reported living as a couple were significantly more likely to report feeling lonely “hardly ever” or “never” (Figure 8).



**Living alone or with others**

People who live alone are at greater risk of feeling lonely more often. Those living alone were significantly more likely to report loneliness “often/always” and “some of the time” whilst those living with others were significantly more likely to report “hardly ever” or “never” feeling lonely (Figure 9).



**Renters report feeling lonely more often**

Renters were significantly more likely to report loneliness “often/always” and “some of the time” than those who own their home. People who own their home were significantly more likely to report “hardly ever” or “never” experiencing loneliness (Figure 10). Homeowner households tend to have [greater financial wealth](https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/wealthingreatbritainwave5/2014to2016) and average annual household income is [greater for those in owner occupied homes than those in rented homes](https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/adhocs/007357incomeandsourceofincomebyhouseholdtenureforallukhouseholds2015to2016). It may not be housing tenure in itself that is associated with how often people feel lonely but financial security.



**Notes**:

1. ‘Homeowner’ includes own outright, buying with a mortgage or loan, and part buy/part rent.

**People who feel they belong to their neighbourhood less strongly report feeling lonely more often**

Those who felt a sense of belonging to their neighbourhood “not very strongly” or “not at all strongly” were significantly more likely than those with a stronger sense of belonging to report experiencing loneliness “often/always”. This suggests those who feel they belong to their neighbourhood less strongly are at greater risk of loneliness (Figure 11).



**People less satisfied with their local area report feeling lonely more often**

Those who reported being “neither satisfied nor dissatisfied” or “fairly/very dissatisfied” with their local area as a place to live were significantly more likely to report feeling lonely “often/always”. Consistent with this, those who reported feeling “very satisfied” with their local area were significantly more likely to report “never” feeling lonely (Figure 12).



# **Which factors independently affect loneliness?**

In order to understand more about how each of these personal characteristics and circumstances contribute to loneliness, logistic regression has been used in the analysis reported in this section. This is an analytical technique that works by focusing on one factor at a time, while holding others constant so the effect of specific characteristics and circumstances can be assessed.

In reality, individuals will have a number of characteristics that could increase or decrease the chances of them feeling lonely and so it can be difficult to identify the underlying causes of loneliness, or alternatively, what keeps people from feeling lonely. Using this technique, it is possible to isolate the different effects of each characteristic or circumstance on reported loneliness.

All factors reported in this section have a statistically significant link with loneliness (that is, we are confident these findings are robust and not just due to random variability in the survey estimates). Regression analysis can identify relationships between factors, however, it cannot tell us about causality.

Of the 34 characteristics and circumstances included in this analysis, 13 were found to have an impact on loneliness, including:

* age
* sex
* marital status
* respondent and partners’ (if applicable) gross income
* disability status (self-reported)
* general health (self-reported)
* number of adults in the household
* caring responsibilities
* whether chat to neighbours more than to just say hello
* feeling as though you belong to a neighbourhood
* satisfaction with local area as a place to live
* the number of years lived in local neighbourhood
* how often meet up in person with family members or friends

When all other factors are held constant, the likelihood of reporting feeling lonely more often tends to decrease with age. The 25 to 34 years, 65 to 74 years and 75 years and over age groups were all significantly less likely to be lonely more often than the 16 to 24 years age group.

Those aged 75 or over are 63% less likely to report loneliness than those aged 16 to 24 years. This pattern was also noted in the [previous section](https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/lonelinesswhatcharacteristicsandcircumstancesareassociatedwithfeelinglonely/2018-04-10#who-is-lonely-more-often) but is more prominent here after holding other factors constant related to both age and loneliness such as becoming a widow or having a disability.

There could be different explanations for this. For example, it’s possible that people become more resilient to loneliness as they get older, possibly through the experience of significant life events and life transitions. On the other hand, some research evidence shows that loneliness is associated with poorer life expectancy. A study conducted in 2015 found that [loneliness is related to an increased mortality risk of 26%](http://journals.sagepub.com/doi/full/10.1177/1745691614568352) (after adjusting for age, gender, socio-economic status and pre-existing health conditions). This might mean that people who are lonelier also live shorter lives and are therefore less likely to be represented among the older population.

Health and disability are strongly related to loneliness. A 2017 study shows that [67% of disabled people have felt lonely in the past year](https://www.scope.org.uk/press-releases/nearly-half-of-disabled-people-chronically-lonely?utm_content=buffer3bece&utm_medium=social&utm_source=twitter&utm_campaign=buffer). The logistic regression shows that those with a long-term health condition or disability were 56% more likely to report loneliness than those without. Similarly, those who describe their general health as “fair”, “bad” and “very bad” were 88% more likely to say they feel lonely than those with “good” and “very good” health. It is, however, important to remember that the relationship between health and loneliness could be reciprocal – with poorer health or disability influencing the experience of loneliness as well as loneliness influencing poorer health and disability. For example, Lund and others (2010) argue that [lonely individuals are also at higher risk of the onset of disability](https://academic.oup.com/ageing/article/39/3/319/41042).

Those who have caring responsibilities were found to be 37% more likely to report loneliness than those who do not. [Research by Carers UK (2015)](https://www.carersuk.org/for-professionals/policy/policy-library?task=download&file=policy_file&id=5100) found that 57% have lost touch with family and friends as a result of their caring role, over a third (36%) feel uncomfortable talking to friends about caring, and 49% have experienced difficulties in their relationship with their partner because of their caring role; all three of these aspects have been highlighted as related to loneliness within this research.

Social connections are an important aspect of loneliness and the results here clearly show that communications with friends, family and neighbours, as well as feelings of belonging to and satisfaction with local area are associated with loneliness. Those who said they meet up with friends and family less often also reported higher levels of loneliness. Those who see friends and family less than once a month or never were 84% more likely to be lonely than those who meet up with friends and family daily. Additionally those who never chat to their neighbours were 43% more likely to feel lonely than those who do.

It is important to remember that logistic regression is good for identifying which specific characteristics and circumstances are most strongly related to loneliness. However, in reality, individuals experience a combination of different characteristics and circumstances in their lives and how these come together may be particularly relevant to the perception of loneliness. We’ve undertaken latent class analysis (LCA) as well to explore this in further detail.

# **Profiles of loneliness**

Latent class analysis (LCA) is a statistical technique used to group individuals with similar patterns of characteristics including reported experience of loneliness.

The LCA results identify sets of characteristics that predict how often people report feeling lonely.

We’ve found four distinct profiles of loneliness: three with sets of characteristics and circumstances found to put people at greater risk of feeling lonely more often and a fourth found to reduce this risk. Indicator variables included in the LCA model were:

* marital status
* general health
* long-term physical or mental health conditions
* living alone or with others
* housing tenure
* age group

Having produced profiles based on these characteristics, these were then analysed in terms of additional variables to find out more about the groups. These variables included in the LCA model were:

* sex
* whether or not in paid work
* living as a couple or not
* [personal well-being scores for life satisfaction, happiness, anxiety and worthwhile](https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingintheukqmi)
* economic status (in employment, unemployed or economically inactive)
* whether or not a long-term limiting illness or disability is reported
* strength of belonging to neighbourhood
* trust in others living in neighbourhood
* level of deprivation in area in which they live based on the [English Index of Multiple Deprivation 2015](https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015) (Lower layer Super Output Area)

**What are the combined characteristics of the most lonely people**?

Three sets of characteristics were found to be associated with greater risk of feeling lonely more often.

**Widowed older homeowners living alone with long-term health conditions**

Older widowed homeowners who live alone and have long-term health conditions were particularly likely to report feeling lonely more frequently. The people in this group tended to be:

* widowed
* in worse general health
* living alone
* homeowners
* aged 65 years or older
* have a long-term physical or mental health condition

Of all individuals in this group 69% reported that they felt lonely “occasionally” or more frequently. This compared with 46% in the sample overall who reported feeling lonely as frequently.

On further examination, people in this group were predominantly:

* female
* not in paid work and economically inactive; given their age likely to be retired
* better-off financially than the sample average; as well as being homeowners, 62% of this group live in the 50% least deprived areas
* in terms of personal well-being scores, mean scores for this group are similar to, though marginally worse than the average for the entire sample

**Unmarried, middle-agers, with long-term health conditions**

At even greater risk of feeling lonely more often were unmarried middle-agers also with long-term health conditions. People in this group were characterised as:

* single (never married), separated, or divorced
* living alone but more likely to be renting than owning their own home
* reporting a long-term physical or mental health condition
* unlikely to describe their general health as “very good” or “good” (and so have “very bad” to “fair” health)
* aged 35 to 64 years

In this group 81% of individuals reported that they felt lonely “occasionally” or more frequently, compared with 46% of the sample overall.

Further examination of this group showed that they tended to be:

* less likely to be in paid work
* more likely to be unemployed or economically inactive
* much more likely to report a long-term illness or disability described as “limiting”
* worse off financially than the sample average; 69% of this group live in the 50% most deprived areas
* in terms of personal well-being scores, mean scores for this group are substantially worse than the means for the overall sample and the other loneliness profile groups

**Younger renters with little trust and sense of belonging to their area**

One younger group were identified as experiencing loneliness more often. Respondents in this group were characterised as:

* single, separated, or divorced
* living with others
* renting
* in “good” or “very good” health without any long-term health conditions or disabilities
* aged 16 to 34 years

In this group 61% of individuals reported that they felt lonely “occasionally” or more frequently, compared with 46% of the sample overall.

Further examination of this group showed that they tended to be:

* likely to be in paid work
* living as a couple (53%) (suggesting that although their marital status indicated being single, just over half were cohabiting)
* without a strong sense of belonging to their neighbourhood; 55% reported feeling that they belonged to their neighbourhood “not very strongly” or “not at all”, compared with 38% in the sample overall
* have little trust of others living in their neighbourhood; only 25% reported feeling that “many” living in their neighbourhood can be trusted, compared with an average of 45%
* worse off financially; as well as being renters, 70% of this group live in the 50% most deprived areas or neighbourhoods

**Comparisons between the loneliness profile groups**

Common to two of the loneliness profile groups was the experience of having a long-term health condition or disability. It is possible that health problems or disability may be factors in the greater frequency of reported loneliness in these groups.

Of the Widowed older homeowners living alone with long-term health conditions group, 65% reported having long-term physical or mental health problems and 53% reported having a long-term illness or disability that was “limiting”. This compares with 33% and 22% in the sample overall respectively. To some extent this may be unsurprising given the higher age of this group. However, for the Unmarried, middle-agers with long-term health conditions group, despite being younger and so at lesser risk of age-related health conditions, a greater proportion reported health problems and described these as “limiting”. A very high proportion (90%) of this group reported having long-term physical or mental health problems and 79% described a long-term illness or disability that was “limiting”.

This analysis, however, emphasises that those in good health and without disability can also be at risk of experiencing loneliness as illustrated by the other often lonely profile group who were younger and in good health without limiting illness or disability.

Our analysis also highlights that people can experience loneliness at any age, though there are different circumstances and characteristics associated with loneliness at different ages.

People who rent rather than owning their homes were prominently represented among two of the loneliness profile groups. Our earlier analysis also found that those who rent their home report loneliness more frequently than those who own their homes. We have not differentiated between renting private or social housing and cannot be sure whether this finding reflects different financial positions among those who rent and own, or whether it may be possibly related to differences in a sense of belonging to their local area. However, the group we’ve described as Widowed older homeowners living alone with long-term health conditions emphasises that those better off financially can also be at risk of loneliness given their other circumstances (for example, widowhood).

**Characteristics of the least lonely profile group**

This analysis so far has focused on the characteristics and circumstances of those most lonely, however, one group were also identified who were least lonely.

**Married homeowners in good health living with others**

The characteristics of this group are:

* living with a partner in a marriage or civil partnership
* better general health
* own their own home
* do not live alone
* are aged 35 years or older (though skewed more towards 65 years and over)

Only 15% of individuals in this group reported that they felt lonely “occasionally” or more frequently, compared with 46% of the sample overall.

This group tended to be:

* male (61%)
* living as a couple (99%)
* with a strong sense of belonging to their neighbourhood; 79% reported that they felt they belonged to their neighbourhood “very” or “fairly strongly”, compared with 62% of the sample overall
* substantially better off in terms of well-being scores than all other loneliness profile groups and the sample means
* in terms of economic activity status, none were unemployed, 15% were in employment and 85% were economically inactive; the data indicate that most were economically inactive because they were retirement age or above

This set of characteristics might be more protective in terms of loneliness.

# **Annexes**

# **Annex 1: Technical report**

# **Introduction**

Using data from the [Community Life Survey August 2016 to March 2017](https://www.gov.uk/government/collections/community-life-survey), bivariate analysis was initially carried out to explore possible associations between a range of individual characteristics and circumstances and self-reported loneliness. This was followed by further, more in-depth analyses to explore the nature and relative strength of these relationships with loneliness. The aim has been to produce in-depth insights to help decision makers target initiatives to alleviate loneliness more effectively.

The research reported here used an iterative research programme involving descriptive analysis followed by logistic regression and finally, latent class analysis (LCA). The logistic regression and LCA analysis approach the exploration of loneliness from two different, but complementary, standpoints. Whilst the logistic regression seeks to isolate single factors that impact on the likelihood of loneliness, LCA seeks to identify combinations of factors that frequently appear together among those who report loneliness. This helps to provide a more holistic picture and highlights that, in practice, it may be a combination of multiple characteristics and circumstances that together shape our experiences and perceptions of loneliness. The annexes provide technical information about how these techniques were applied.

* 1. **The Community Life Survey 2016 to 2017 data**

The research relied on data from the annual Community Life Survey (CLS), a nationally representative household survey of adults (aged 16 and over) in England. The CLS 2016 to 2017 dataset contains data for 10,256 adults for the period August 2016 to March 2017.

The CLS 2016 to 2017 dataset was selected for analysis because the survey asked respondents about their frequency of loneliness. The survey also solicited information about the respondents’ socio-demographic characteristics, behaviours, attitudes, community engagement and circumstances, which were used as explanatory variables.

**Loneliness: the outcome variable**

Central to the analysis was the question included in the CLS 2016 to 2017, which asked respondents: How often do you feel lonely?

1. Often/always
2. Some of the time
3. Occasionally
4. Hardly ever
5. Never

For the purposes of this report this is referred to as “the loneliness question”.

**(Re)coding variables for analysis**

**Dichotomising loneliness**

A binary version of the loneliness variable was used for the logistic regression and LCA. Responses of “often/always”, “some of the time”, and “occasionally” were collapsed into a single category of “more often lonely”, and those of “hardly ever” or “never” into another of “hardly ever or never lonely”. Whilst dichotomising the outcome variable in this way obscures some differentiation between frequency categories of reported loneliness, it was necessary for the logistic regression and LCA techniques. Reasons for recoding loneliness in this way are detailed in this section.

There is a relatively small sample size. The CLS 2016 to 2017 dataset contains responses from 10,256 individuals and, of these, 10,057 cases have valid data for the loneliness question. For a case to be included in the LCA model there must be valid data for every variable included in the model. With inclusions of each additional variable there is greater likelihood that any given case will become ineligible due to missing data and so be excluded from the model. In the final logistic model and LCA specification (see section 3 and 4 respectively), the sample size was reduced to 6,414 and 6,149 respectively because of missing data.

For reasons of statistical quality, it was decided that explanatory variables should, ideally, be tabulated with the binary loneliness variable so that wherever possible all (unweighted) cell counts are at least 100. This “100 minimum cell count” rule was relatively arbitrary but it was decided that some sort of minimum count was needed. This rule was achieved in all variables except for economic activity where, due to relatively small numbers of unemployed in the sample, 60 (unweighted) cases reported unemployment and that they experienced loneliness “hardly ever” or “never”.

Whilst it was necessary to recode variables to have fewer categories, ideally recoding should preserve the underlying distribution whilst having fewer categories ([Strait et al., 1996](https://watermark.silverchair.com/45-1-67.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAcUwggHBBgkqhkiG9w0BBwagggGyMIIBrgIBADCCAacGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQMRyvd6PNkRkWR2NrFAgEQgIIBeMu6u9VRq6QSA1T63IqgyyogHGbPlMYJj-jtTmS9X1SHNzUnAiYqCkCxKuPrlTuQ4POGnp5bJnPWySNtp2hJVzxnS-ar4Sr1veUAIcMHEBumo6RUoB1E5VinAZjBxmVMDHD7DTFORY2L_ldyKu0mqQsuo-_-LqhUS2-2agutIHdlrTAc0EEegt5-5_9o-VCCbgi69L-57LYgzcaCtyDVvKu8cyk4W77p-1agwEeEPF1uoRgopFYRAlYj2FFZy7p_vB0__8QgFPsk_Wy1fvk28t1AUIugjdMx8cMTCDF0oT37YX4zmBYHbBPxXSKPTpxqtJ_eOevxzDzg1S6wpzwoA0HVmMuePjynqXjl0hnz_upJKgzi-sqpFCQ0kp_hZjoU5V-7wcArrRN_izrorwuXdyq3C9n8Lp271wBRObvIV82FsS7byTkUJ6LAvPh9w7YxyUCF4Bx1H3KWK6TVaJM_1AooAve9kYbbLukgcYGbqI2DvRurWaNlTCI)). The distribution of responses to the loneliness question is shown in Figure 1.



This shows that the frequency of loneliness is skewed towards the “hardly ever” and “never” end of the response scale. By dichotomising the loneliness variable as described previously, categories had broadly similar frequencies of respondents thereby broadly preserving the distribution of the original variable: 4,841 were “more often lonely” and 5,216 as “hardly ever or never lonely”. With a larger sample size, it may have been possible to include more categories of loneliness thereby aiding greater differentiation in terms of loneliness frequency.

Another reason is consistency between coding for the logistic regression and LCA. As the LCA (for the reasons described previously) required a binary version of the loneliness variable, for consistency of results it made sense to apply a form of logistic regression that uses binary coding. Additionally, while it is possible to conduct multinomial logistic regression with multiple categorical outcomes, logistic regression with binary outcomes (for example, “lonely” compared with “not lonely”) is also easier to interpret and explain.

**Recoding (and deriving) explanatory or independent variables**

In many instances, independent or explanatory variables needed further preparation before inclusion in the models.

As noted earlier, it is better to preserve the original distribution of variables as much as possible when recoding for LCA and this was taken into consideration when recoding explanatory variables. Also, (as noted earlier) missing data is problematic. Therefore, variables that had more than 3,000 missing cases were excluded.

Small cell counts can produce poor quality analysis. As noted earlier, to ensure that when each explanatory variable was tabulated with the loneliness variable there was a minimum cell count of 100, categories were collapsed and, where appropriate, some categories were recoded as missing, thereby removing those cases from analysis. After recoding, and as already noted, only economic status broke this rule due to a relatively small number of unemployed people in the sample.

Greater importance, though, was given to producing recodes that were useful for meaningful interpretation – categories were only collapsed where the new category made sense. For example, it would not have been meaningful to collapse unemployed people into any other economic category.

**Missing data and bias**

As noted, cases with missing data for variables included in the LCA model are excluded from analysis. Missing data can produce biased estimates and invalid conclusions, particularly if data are not “missing at random” or, in other words, if there is some (unknown) patterning to that “missingness” ([Graham, 2009](https://www.ncbi.nlm.nih.gov/pubmed/18652544)).

We have not examined missing data in our analysis and we do not know if, or to what extent, some people with particular characteristics may fail to provide responses more than people with different characteristics. We did not use any techniques for dealing with missing data (for example, imputation). Consequently, we cannot know if or how the patterning of missing data impacted on our findings.

* 1. **Logistic regression**

Logistic regression analysis allows for the relationship between an explanatory variable and the outcome variable to be examined, whilst at the same time taking into consideration other explanatory variables that influence the outcome. Logistic regression is used as it is suitable when looking at categorical outcomes (which is the form taken by most of the Community Life Survey (CLS) variables). While it is possible to conduct multinomial logistic regression with multiple categorical outcomes, logistic regression with binary outcomes (for example, “lonely” compared with “not lonely”) was chosen. This was chosen to increase ease of understanding (with the predicted outcomes being either “lonely” or “not lonely”); and for consistency with the LCA.

**Procedure**

This analysis has been carried out in SAS 9.3. All variables have been treated as categorical variables. The sample size for the logistic regression analysis is 6,414. Backwards logistic regression was used to create the final model. The contribution of each variable is assessed by looking at the significance value of the t-test for each predictor. If there is at least one non-significant variable, the variable with the highest p-value is removed from the model. This procedure is repeated, until the all the remaining variables are significant at the 0.05 level.

There are multiple ways in which variables could be entered in to the model. Forward, backwards and stepwise models were tried and it was found that most of the variables were the same in each case. The backward logistic regression method was used for the final model as it produced a model with the lowest Akaike Information Criterion (AIC); additionally, forward approaches often allow for important variables to be missed due to other variables being entered in to the model first (“suppressor effects”).

**Multicollinearity**

Many of the variables collected in the Community Life Survey are correlated with one another. Multicollinearity (also known as collinearity) is where one or more explanatory variables in a regression model are highly correlated such that they linearly predict each other with a high degree of accuracy. However, an important assumption of multivariate regression is that explanatory variables are not too highly correlated with one another. Too high a degree of correlation between predictor variables in a regression model can affect the stability and interpretation of the regression estimates.

In the final model, there were a few variables that were correlated, however, their absolute Pearson’s Correlation value was less than 0.5 and the model performs better including these variables and so they have remained in the model. These are disability and health (Pearson’s correlation figure of negative 0.46463), and chatting to neighbours, belonging to the neighbourhood and satisfaction with the local area (Pearson’s correlation figure of 0.31267 for chatting to neighbours and belonging to the neighbourhood, 0.16419 for satisfaction with the local area and chatting to neighbours, and 0.39001 for belonging to the neighbourhood and satisfaction with the local area).

**Goodness of fit**

Goodness of fit describes how well a model fits the data from which it is generated. It can be used to assed how well the data that the model predicts and corresponds to the data that have been collected. There are various measurements used to assess the model fit. The first two, AIC and Schwarz Criterion (SC) are deviants of negative two times the log-likelihood (-2 Log L). AIC and SC penalize the log-likelihood by the number of predictors in the model. AIC and SC are used for the comparison of non-nested models on the same sample. Ultimately, the model with the smallest AIC and SC are considered the best, although the AIC and SC value itself is not meaningful.

The Likelihood Ratio (LR) Chi-Square test, the Score Chi-Square Test and the Wald Chi-Square Test all test that at least one of the predictors’ regression coefficient is not equal to zero in the model. The Residual Chi-Square Test shows the Chi-Square test statistic, the degrees of freedom (DF) and the associated p-value (PR>ChiSq) corresponding to the specific test that all of the predictors are simultaneously equal to zero. A small p-value from all three tests leads to the conclusion that at least one of the regression coefficients in the model is not equal to zero.

**Interaction effects**

Interactions can be used to test for the joint effect of two or more predictor variables on an outcome variable. It allows us to explore how the relationships between dependent and independent variables differ by context. Some interactions were identified as being significant, however, there is no prior evidence to support the link with loneliness. Some of the interactions appeared to be counter intuitive and did not have a large improvement to the model in terms of improving the AIC. Additionally, adding an interaction term to a model drastically changed the interpretation of all of the coefficients in the model. It was decided, for the purpose of this analysis, to remove interactions for the benefit of identifying individual impacts of each variable.

**Causality**

Regression analysis can identify relationships between factors; however, it cannot tell us about causality. While, for some factors, causality is fairly clear based on prior knowledge (for example, loneliness does not cause someone to become widowed, however, becoming widowed can cause loneliness), for others the relationship between cause and effect is more blurred (for example, ill health can cause loneliness, but also loneliness can cause ill health). Therefore, where prior knowledge does not make the direction of causality clear it’s important to note that causality can operate in either direction (or both).

**Weighting**

The results of the Community Life Survey are weighted to compensate for unequal selection probabilities and differential non-response (that is, to ensure that the age and sex distribution of the final dataset matches that of the population of England). Our regression models take the weights into account.

**Interpretation of the results**

The odds ratio is the usual output from logistic regression. The odds ratio for each variable in the model is obtained by exponentiating the estimate. The odds ratio can be interpreted as follows: for a one-unit change in the predictor variable, the odds ratio for a positive outcome is expected to change by the respective coefficient, given the other variables in the model are held constant.

The 95% Wald Confidence Limits are provided for each odds ratio. For a given predictor variable with a level of 95% confidence, that upon repeated trials, 95% of the confidence interval (CIs) would include the “true” population odds ratio. The CI is equivalent to the Chi-Square test statistic: if the CI includes one, the null hypothesis that a particular regression coefficient is equal to zero and the odds ratio is equal to one, given the other predictors are in the model would fail to be rejected. An advantage of a CI is that it is illustrative; it provides information on where the “true” parameter may lie and the precision of the point estimate for the odds ratio.

# **Latent Class Analysis**

Latent class analysis (LCA) is a statistical technique used to identify sub-groups within a population. Applied to survey data, LCA classifies individuals into groups or “types” based on patterns of characteristics represented as categorical variables. LCA was used to group individuals with similar patterns of characteristics including reported experience of loneliness. By employing LCA as reported here, combinations of characteristics that “go with” experience of loneliness are revealed.

Some combinations were found to characterise groups that were more frequently lonely (these factors may be risky in terms of loneliness) whilst other characteristics were found to characterise groups that were less frequently (or never) lonely (these factors may be more protective against loneliness). It is reasonable to think of these characteristics in terms of profiles. Using LCA in this way can aid the identification of groups in the general population who exhibit combinations of characteristics that put them at greater risk of loneliness and others with characteristics more protective in terms of loneliness.

**LCA approach taken**

The loneliness variable was included within the model along with other variables and then, by adding and taking away variables one-by-one, the aim was to produce a model with good separation (particularly on the loneliness variable). Another method would have been to split our dataset in terms of responses to the loneliness question prior to developing a LCA model. For example, a subset of the data could have been taken to include only those who reported feeling lonely “often/always” and then tested some variables for good separation – this may have produced various groups with different similar characteristics all of which were most frequently lonely. Similarly, a subset of data could have included only those cases in the LCA model who report being less lonely (for example, never).

However, these approaches were not taken for two main reasons. Firstly, use of the full dataset (rather than a subset) allows for better comparisons between people with different characteristics across all variables including the loneliness variable. Secondly, the relatively small sample size would have been reduced further leading to poorer quality results.

**Selection of explanatory variables for the final LCA specification**

The logistic regression highlighted characteristics that significantly increase or decrease likelihood of loneliness if all other factors are held constant. As a starting point in building the LCA specification, these were used to build LCA models. Through trial and error, adding and taking away one variable at a time and re-running the algorithm, a model specification was produced using the variables pertaining to the following:

**Loneliness frequency:**

* 1 = Often/always, Some of the time; Occasionally
* 2 = Hardly ever; Never

**Marital status:**

* 1 = single, that is, never married and never registered in a same-sex civil partnership; Separated/divorced
* 2 = Living with partner in a marriage or civil partnership (and not separated)
* 3 = widowed

**General health:**

* 1= Very good or good
* 2 = Fair
* 3 = Very bad or bad

**Housing Tenure:**

* 1 = Own outright/buying with mortgage/loan/part buy part rent
* 2 = Renting

**Presence or absence of a physical or mental health condition/illness lasting or expected to last 12 months or more:**

* 1 = Yes
* 2 = No

**Lives alone or does not live alone:**

* 1 = Lives alone
* 2 = Does not live alone

**Age grouped into three categories:**

* 16 to 34
* 35 to 64
* 65 and over

**Identifying lonely groups or profiles**

LCA is undertaken to produce groups of individuals with different characteristics so that individuals within groups are more similar to each other while, at the same time, distinct from other groups. Table 1 presents figures for the final LCA model.

A model with better separation has less equal distribution between each group in terms of variable categories – in general, values approaching 100% indicate clearer delineation between groups ([Celeux and Soromenho, 1996](https://link.springer.com/article/10.1007/BF01246098)). As our focus was loneliness, it was important that our LCA output showed good separation in terms of the loneliness variable. For example, in Table 1 Group C shows the best separation of all with 85% of individuals reporting “hardly ever” or “never” feeling lonely and 15% who reported feeling lonely “often/always”, “some of the time” or “occasionally”. Of course, a more useful model also provides good separation in terms of other variables included – unequal distributions and deviations from the mean are particularly worth noting because this suggests characteristics that differ from the average and/or other groups.

Based on our data, a deviation from the mean of 15% was chosen for identifying lonely and non-lonely groups. As shown in Table 1, there are four groups that fulfil this criterion: groups A, C, D and E. In the main article, we only report on these groups because these had distributions of loneliness most different from the mean. For transparency, Table 1 presents all seven groups produced by the LCA model. For the raw LCA data, see Appendix 2.

In the main article, we refer to:

* Group A as the Widowed older homeowners living alone with long-term health conditions group
* Group C as the Married homeowners in good health living with others group
* Group D as the Unmarried, middle-agers, with long-term health conditions group
* Group E as the Younger renters with little trust and sense of belonging to their group

**Optimal number of groups**

The LCA process involves running the algorithm with different numbers of groups specified. The analyst first specifies one group, then two groups, then three and so on. With each run a goodness of fit statistic, the Bayes Information Criterion (BIC), is produced. In exploratory LCA, the BIC coefficient is used to identify the optimal number of classes ([Lin and Dayton 1997](http://journals.sagepub.com/doi/abs/10.3102/10769986022003249?journalCode=jebb)) and in line with this, the number of groups with the lowest BIC coefficient was chosen as the best model. A model with seven classes was identified to be best – see Appendix 2 for the BIC coefficients of models with one through to eight classes.

In Table 1, groups A, C, D and E show good separation in terms of loneliness. These groups have loneliness responses that differ from the mean proportion of the sample by at least 15% in terms of loneliness. Looking at the whole sample, 46% of people fall into the “more often lonely” category whilst in group A, for example, 69% of people fall into the “more often lonely” category – a much higher proportion than the sample’s average.

|  |
| --- |
| **Table 1: Groups and characteristics included the latent class analysis model** |
| England |  |  |  |  |  |  |  | Counts  |
|  |  | Group % |
|  | Category | A | B | C | D | E | F | G | Mean |
| Loneliness | More often lonely | 69 | 38 | 15 | 81 | 61 | 57 | 57 | 46 |
| Hardly ever or never lonely | 31 | 62 | 85 | 19 | 39 | 43 | 43 | 54 |
| Marital status | Single, separated or divorced | 26 | 16 | 2 | 92 | 76 | 95 | 2 | 35 |
| Married or civil partnership | 4 | 83 | 98 | 3 | 24 | 2 | 96 | 61 |
| Widowed | 70 | 1 | 1 | 4 | 0 | 3 | 2 | 4 |
| General health | Very good or good | 48 | 91 | 86 | 16 | 89 | 93 | 7 | 74 |
| Fair | 41 | 9 | 14 | 51 | 10 | 7 | 66 | 20 |
| Very bad or bad | 11 | 0 | 0 | 33 | 1 | 0 | 28 | 6 |
| Tenure | Homeowner | 88 | 86 | 97 | 40 | 22 | 74 | 80 | 72 |
| Renting | 12 | 14 | 3 | 60 | 78 | 26 | 20 | 28 |
| Long-term health condition | Yes | 65 | 13 | 37 | 90 | 11 | 15 | 92 | 33 |
| No | 35 | 87 | 63 | 10 | 89 | 85 | 8 | 67 |
| Lives alone | Yes | 88 | 0 | 2 | 54 | 3 | 90 | 0 | 16 |
| No | 12 | 100 | 98 | 46 | 97 | 10 | 100 | 84 |
| Age group | 16-34 | 0 | 15 | 0 | 18 | 80 | 15 | 3 | 22 |
| 35-64 | 4 | 81 | 36 | 71 | 20 | 70 | 56 | 55 |
| 65+ | 96 | 4 | 64 | 10 | 0 | 16 | 41 | 23 |
| Source: 'Community Life Survey 2016 to 2017', Department for Digital, Cultural, Media & Sport |
| Notes: |
| 1. Single, separated or divorced includes those who have never married or registered in a same-sex civil partnership, and those who are divorced or separated (but may still be legally married). |
| 2. Homeowner includes those own their home outright, are buying with a mortgage/loan, and part buy, part rent.  |
| 3. Renting includes those renting |
| 4. Those who reported living rent free or occupy in any other way are coded as missing due to small cell counts.  |
| 5. A long-term health condition means that respondents reported having a "physical or mental health conditions or illnesses lasting or expected to last for 12 months or more" |

The groups identified are dependent on the variables included in the LCA model. Had other variables been included then the groups produced would have been different. Unlike some other statistical techniques (for example, logistic regression), variable selection is less automated by the algorithm and more dependent on the choices of the analyst. The absence or presence of a single variable can change whether good separation is achieved or not, and/or how any groups are found optimal. There are practically countless combinations of variables and codes and it is not possible to test them all.

**Additional descriptive statistics**

In the final LCA algorithm, only the variables and categories as shown in Table 1 were included. In general, with additional variables included in the model there was poorer separation in terms of loneliness across clusters. Good separation in terms of loneliness was the main focus. However, when fewer variables were included, the LCA model became less informative because there was less differentiation in terms of other characteristics, simply because these variables were not included in the model. It is therefore a balance between producing good separation on loneliness and with including more variables that can contribute to, and can be used to describe, the groups. Table 2 presents the characteristics of all seven groups in terms of additional descriptive statistics.

|  |
| --- |
| Table 2: Groups and additional characteristics |
| England |
| Group |
|  |  | A |  | B |  | C |  | D |  | E |  | F |  | G |  | Average |
| Median age (years) |  | 73 |  | 46 |  | 70 |  | 49 |  | 28 |  | 53 |  | 62 |  | 49 |
| Sex (%) | Male | 33 |  | 49 |  | 61 |  | 44 |  | 40 |  | 47 |  | 53 |  | 48 |
|  | Female | 67 |  | 51 |  | 39 |  | 56 |  | 60 |  | 53 |  | 47 |  | 52 |
| Paid job (%) | Yes | 9 |  | 82 |  | 12 |  | 48 |  | 76 |  | 73 |  | 35 |  | 61 |
|  | No | 91 |  | 18 |  | 88 |  | 52 |  | 24 |  | 27 |  | 65 |  | 38 |
| Living as a couple? (%) | Yes | 5 |  | 94 |  | 99 |  | 15 |  | 53 |  | 2 |  | 98 |  | 71 |
|  | No | 95 |  | 6 |  | 1 |  | 85 |  | 47 |  | 98 |  | 2 |  | 29 |
| Life satisfaction (mean) |  | 6.87 |  | 7.37 |  | 8.26 |  | 5.29 |  | 6.97 |  | 6.98 |  | 6.36 |  | 7.1 |
| Happiness (mean) |  | 7.07 |  | 7.37 |  | 8.24 |  | 5.13 |  | 7 |  | 7.07 |  | 6.29 |  | 7.1 |
| Anxiety (mean) |  | 3.28 |  | 3.3 |  | 2.25 |  | 5.06 |  | 3.7 |  | 3.19 |  | 3.98 |  | 3.44 |
| Worthwhile (mean) |  | 6.94 |  | 7.61 |  | 8.21 |  | 5.55 |  | 7.11 |  | 7.2 |  | 6.59 |  | 7.28 |
| Economic status (%) | Employed | 10 |  | 82 |  | 15 |  | 49 |  | 76 |  | 75 |  | 37 |  | 62 |
|  | Unemployed | 0 |  | 2 |  | 0 |  | 5 |  | 4 |  | 3 |  | 2 |  | 2 |
|  | Inactive | 90 |  | 16 |  | 85 |  | 46 |  | 20 |  | 22 |  | 61 |  | 35 |
| Limiting long-term health condition (%) | Yes | 53 |  | 7 |  | 16 |  | 79 |  | 5 |  | 7 |  | 79 |  | 22 |
|  | No | 47 |  | 93 |  | 84 |  | 21 |  | 95 |  | 93 |  | 21 |  | 78 |
| Neighbourhood strength of belonging (%) | More strongly | 67 |  | 67 |  | 79 |  | 47 |  | 45 |  | 53 |  | 66 |  | 62 |
|  | Less strongly | 33 |  | 33 |  | 21 |  | 53 |  | 55 |  | 47 |  | 34 |  | 38 |
| Trust in peopleliving in neighbourhood (%) | “Many people can be trusted” | 61 |  | 50 |  | 68 |  | 27 |  | 25 |  | 41 |  | 41 |  | 45 |
|  | “Some can be trusted” | 23 |  | 32 |  | 22 |  | 33 |  | 37 |  | 36 |  | 34 |  | 32 |
|  | “A few can be trusted” | 16 |  | 16 |  | 9 |  | 29 |  | 31 |  | 19 |  | 22 |  | 19 |
|  | “None can be trusted” | 0 |  | 2 |  | 0 |  | 11 |  | 7 |  | 4 |  | 2 |  | 3 |
| English Index of Multiple Deprivation 2015 (LSOA) (%) | Bottom 50% | 38 |  | 43 |  | 27 |  | 69 |  | 71 |  | 53 |  | 44 |  | 49 |
|  | Top 50% | 62 |  | 57 |  | 73 |  | 31 |  | 29 |  | 47 |  | 56 |  | 51 |
| Source: 'Community Life Survey 2016 to 2017', Department for Digital, Cultural, Media & Sport |

# **Annex 2: Logistic regression – statistical explanations and tables**

Initial list of variables considered:

* Mode of Interview
* Age group
* Sex
* Ethnicity
* Relationship status
* Income
* Urban or rural classification
* Region
* Housing tenure
* Disability
* General health
* Education
* Digital skills
* Employment status
* Number of adults
* Number of children
* Volunteering
* Caring responsibilities
* Agree people in neighbourhood pull together
* Whether chat to neighbours more than just to say Hello
* Trust people in neighbourhood
* Belong to neighbourhood
* Religion (even if not practicing)
* Satisfaction with local area as a place to live
* Has area got better or worse in last two years
* Years lived in neighbourhood
* Number of services and amenities in local area
* Index of Multiple Deprivation
* National Statistics Socio-economic Classification (NS-SEC)
* This local area is a place where people from different backgrounds get on well together?
* How often meet up in person with family members or friends
* How often speak on the phone or video or audio call via the internet with family members or friends
* How often email or write to family members or friends
* How often exchange text messages or instant messages with family members or friends

Variables removed as not being significant predictors on their own:

* Religion was removed as it is not correlated with loneliness using the Pearson product-moment correlations. The correlations range from negative 1 to positive 1, and the Pearson product-moment correlation between religion and loneliness is 0.00148 (p equals 0.8827).

Variables removed as not being significant predictors when part of a regression model:

* Mode of interview
* Ethnicity
* Urban or rural classification
* Region
* Housing tenure
* Education
* Digital skills
* Employment status
* Number of children
* Volunteering
* Agree people in neighbourhood pull together
* Trust people in neighbourhood
* Has area got better or worse in last two years
* Number of services and amenities in local area
* Index of Multiple Deprivation
* NS-SEC
* This local area is a place where people from different backgrounds get on well together?
* How often speak on the phone or video or audio call via the internet with family members or friends
* How often email or write to family members or friends
* How often exchange text messages or instant messages with family members or friends

**Final logistic regression model**

|  |
| --- |
| **Logistic Regression output - Response Profile** |
| Ordered Value | Lonely |  | Total Frequency |  | Total Weight |
| 1 | 0 |  | 3339 |  | 16194897 |
| 2 | 1 |  | 3075 |  | 13721262 |
| Source: 'Community Life Survey 2016 to 2017', Department for Digital, Cultural, Media & Sport |

|  |
| --- |
| **Logistic Regression output: Testing Global Null Hypothesis: BETA=0** |
| Test | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio | 5333515.28 | 37 | <.0001 |
| Score | 4916532.05 | 37 | <.0001 |
| Wald | 658.0485 | 37 | <.0001 |
| Source: 'Community Life Survey 2016 to 2017', Department for Digital, Cultural, Media & Sport |
|  |  |  |  |

|  |
| --- |
| **Logistic Regression output: Type 3 Analysis of Effects** |
| Effect | DF |  | Wald Chi-Square |  | Pr > ChiSq |
| Rage9\_recode | 6 |  | 52.1036 |  | <.0001 |
| Sex | 1 |  | 48.9797 |  | <.0001 |
| MarStatg2 | 3 |  | 46.7504 |  | <.0001 |
| ZIncomhh1 | 2 |  | 16.7081 |  | 0.0002 |
| dill2 | 1 |  | 27.6347 |  | <.0001 |
| ghealth2 | 1 |  | 48.9124 |  | <.0001 |
| nadults | 4 |  | 46.9456 |  | <.0001 |
| RCare | 1 |  | 10.643 |  | 0.0011 |
| chat2neigh | 1 |  | 9.0925 |  | 0.0026 |
| SBeNeigh | 3 |  | 17.8962 |  | 0.0005 |
| Slocsat | 4 |  | 31.3641 |  | <.0001 |
| yearsLived | 5 |  | 17.1413 |  | 0.0042 |
| meetupnew | 5 |  | 33.3145 |  | <.0001 |
| Source: 'Community Life Survey 2016 to 2017', Department for Digital, Cultural, Media & Sport |

|  |
| --- |
| **Logistic Regression output: Analysis of Maximum Likelihood Estimates** |
| Parameter |  | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
| Intercept |  | 1 | 0.4182 | 0.1081 | 14.9774 | 0.0001 |
| Rage9\_recode | 2 | 1 | 0.418 | 0.0921 | 20.5841 | <.0001 |
| Rage9\_recode | 3 | 1 | 0.0972 | 0.0817 | 1.4147 | 0.2343 |
| Rage9\_recode | 4 | 1 | 0.0374 | 0.0837 | 0.1994 | 0.6552 |
| Rage9\_recode | 5 | 1 | 0.0514 | 0.0874 | 0.3457 | 0.5566 |
| Rage9\_recode | 6 | 1 | -0.5309 | 0.0912 | 33.8912 | <.0001 |
| Rage9\_recode | 7 | 1 | -0.5271 | 0.1412 | 13.9349 | 0.0002 |
| Sex | Female | 1 | 0.2425 | 0.0347 | 48.9797 | <.0001 |
| MarStatg2 | 1 | 1 | -0.4894 | 0.0772 | 40.1797 | <.0001 |
| MarStatg2 | 2 | 1 | -0.1122 | 0.0985 | 1.2971 | 0.2547 |
| MarStatg2 | 3 | 1 | 0.7944 | 0.14 | 32.1764 | <.0001 |
| ZIncomhh1 | 1 | 1 | -0.0865 | 0.0465 | 3.4509 | 0.0632 |
| ZIncomhh1 | 2 | 1 | -0.161 | 0.0612 | 6.9128 | 0.0086 |
| dill2 | 0 | 1 | 0.2224 | 0.0423 | 27.6347 | <.0001 |
| ghealth2 |  | 1 | 0.6291 | 0.0899 | 48.9124 | <.0001 |
| nadults | 2 | 1 | -0.1246 | 0.0711 | 3.0671 | 0.0799 |
| nadults | 3 | 1 | -0.1122 | 0.0868 | 1.6695 | 0.1963 |
| nadults | 4 | 1 | -0.074 | 0.0911 | 0.6603 | 0.4165 |
| nadults | 5 | 1 | -0.3339 | 0.1736 | 3.6988 | 0.0545 |
| RCare | Yes | 1 | 0.1573 | 0.0482 | 10.643 | 0.0011 |
| chat2neigh | 2 | 1 | 0.1798 | 0.0596 | 9.0925 | 0.0026 |
| SBeNeigh | Fairly strongly | 1 | 0.0165 | 0.0569 | 0.0842 | 0.7717 |
| SBeNeigh | Not at all strongly | 1 | 0.1541 | 0.0963 | 2.5626 | 0.1094 |
| SBeNeigh | Not very strongly | 1 | 0.1447 | 0.0598 | 5.8572 | 0.0155 |
| Slocsat | Fairly dissatisfied | 1 | 0.1061 | 0.1213 | 0.7644 | 0.382 |
| Slocsat | Fairly satisfied | 1 | -0.0654 | 0.0724 | 0.815 | 0.3667 |
| Slocsat | Neither satisfied nor dissatisfied | 1 | 0.2445 | 0.0957 | 6.5244 | 0.0106 |
| Slocsat | Very dissatisfied | 1 | 0.0968 | 0.1885 | 0.264 | 0.6074 |
| yearsLived | 2 | 1 | 0.0255 | 0.0843 | 0.0917 | 0.762 |
| yearsLived | 3 | 1 | -0.0169 | 0.0909 | 0.0345 | 0.8527 |
| yearsLived | 4 | 1 | -0.1321 | 0.0898 | 2.1642 | 0.1413 |
| yearsLived | 5 | 1 | 0.1558 | 0.0926 | 2.8314 | 0.0924 |
| yearsLived | 6 | 1 | -0.2356 | 0.0924 | 6.4985 | 0.0108 |
| meetupnew | 1 | 1 | -0.2097 | 0.0636 | 10.8763 | 0.001 |
| meetupnew | 2 | 1 | -0.1702 | 0.0677 | 6.3285 | 0.0119 |
| meetupnew | 3 | 1 | -0.0139 | 0.0923 | 0.0227 | 0.8802 |
| meetupnew | 4 | 1 | 0.1778 | 0.1005 | 3.1281 | 0.077 |
|  |  |  |  |  |  |  |
| **Logistic Regression Output: Odds Ratio Estimates** |
| Effect | Point Estimate |  | 95% Wald  |  | Confidence Limits |
| Rage9\_recode 2.00 vs 1.00 | 0.964 |  | 0.72 |  | 1.292 |
| Rage9\_recode 3.00 vs 1.00 | 0.7 |  | 0.516 |  | 0.95 |
| Rage9\_recode 4.00 vs 1.00 | 0.659 |  | 0.48 |  | 0.905 |
| Rage9\_recode 5.00 vs 1.00 | 0.668 |  | 0.476 |  | 0.939 |
| Rage9\_recode 6.00 vs 1.00 | 0.373 |  | 0.262 |  | 0.533 |
| Rage9\_recode 7.00 vs 1.00 | 0.375 |  | 0.242 |  | 0.58 |
| Sex Female vs Male | 1.624 |  | 1.418 |  | 1.861 |
| MarStatg2 1.00 vs 0.00 | 0.743 |  | 0.61 |  | 0.906 |
| MarStatg2 2.00 vs 0.00 | 1.084 |  | 0.821 |  | 1.431 |
| MarStatg2 3.00 vs 0.00 | 2.683 |  | 1.814 |  | 3.969 |
| ZIncomhh1 1.00 vs 0.00 | 0.716 |  | 0.601 |  | 0.853 |
| ZIncomhh1 2.00 vs 0.00 | 0.665 |  | 0.533 |  | 0.829 |
| dill2 0.00 vs 1.00 | 1.56 |  | 1.322 |  | 1.841 |
| ghealth2 | 1.876 |  | 1.573 |  | 2.237 |
| nadults 2.00 vs 1.00 | 0.463 |  | 0.368 |  | 0.584 |
| nadults 3.00 vs 1.00 | 0.469 |  | 0.358 |  | 0.615 |
| nadults 4.00 vs 1.00 | 0.487 |  | 0.367 |  | 0.648 |
| nadults 5.00 vs 1.00 | 0.376 |  | 0.235 |  | 0.602 |
| RCare Yes vs No | 1.37 |  | 1.134 |  | 1.655 |
| chat2neigh 2.00 vs 1.00 | 1.433 |  | 1.134 |  | 1.81 |
| SBeNeigh Fairly strongly vs Very strongly | 1.393 |  | 1.14 |  | 1.703 |
| SBeNeigh Not at all strongly vs Very strongly | 1.599 |  | 1.18 |  | 2.166 |
| SBeNeigh Not very strongly vs Very strongly | 1.584 |  | 1.273 |  | 1.971 |
| Slocsat Fairly dissatisfied vs Very satisfied | 1.629 |  | 1.196 |  | 2.218 |
| Slocsat Fairly satisfied vs Very satisfied | 1.372 |  | 1.169 |  | 1.611 |
| Slocsat Neither satisfied nor dissatisfied vs Very satisfied | 1.871 |  | 1.476 |  | 2.372 |
| Slocsat Very dissatisfied vs Very satisfied | 1.614 |  | 0.999 |  | 2.608 |
| yearsLived 2.00 vs 1.00 | 0.837 |  | 0.676 |  | 1.037 |
| yearsLived 3.00 vs 1.00 | 0.802 |  | 0.632 |  | 1.019 |
| yearsLived 4.00 vs 1.00 | 0.715 |  | 0.564 |  | 0.907 |
| yearsLived 5.00 vs 1.00 | 0.954 |  | 0.744 |  | 1.223 |
| yearsLived 6.00 vs 1.00 | 0.645 |  | 0.501 |  | 0.83 |
| meetupnew 1 vs 0 | 0.987 |  | 0.789 |  | 1.236 |
| meetupnew 2 vs 0 | 1.027 |  | 0.813 |  | 1.298 |
| meetupnew 3 vs 0 | 1.201 |  | 0.907 |  | 1.59 |
| meetupnew 4 vs 0 | 1.455 |  | 1.082 |  | 1.955 |

# **Annex 3: Latent Class Analysis**

|  |
| --- |
| **Bayes Information Criterion coefficients for models with 1 through to 8 classes** |
| Number of classes | Bayes Information Criterion (BIC) |
| 1 | 60013.12 |
| 2 | 57058.24 |
| 3 | 55265.58 |
| 4 | 54367.77 |
| 5 | 53991.97 |
| 6 | 53708.09 |
| 7 | 53609.42 |
| 8 | 53617.31 |
| Source: 'Community Life Survey 2016 to 2017', Department for Digital, Cultural, Media & Sport |